

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

## OM protein - protein search, using sw model

Run on:

March 1, 2001, 15:47:07 ; Search time 210.42 Seconds

(without alignments)  
6.988 Million cell updates/sec

Title: US-09-331-631A-3\_COPY\_74\_116

Perfect score: 250

Sequence: 1 NQDPQTDCQQCORRCRQEE..... RQQQYCRRCKBICEEBEEY 43

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_36:\*

1: /SIDS1/gcdata/geneseq/geneseq/geneseq/geneseq/AA1980.DAT:\*

2: /SIDS1/gcdata/geneseq/geneseq/geneseq/geneseq/AA1981.DAT:\*

3: /SIDS1/gcdata/geneseq/geneseq/geneseq/geneseq/AA1982.DAT:\*

4: /SIDS1/gcdata/geneseq/geneseq/geneseq/geneseq/AA1983.DAT:\*

5: /SIDS1/gcdata/geneseq/geneseq/geneseq/geneseq/AA1984.DAT:\*

6: /SIDS1/gcdata/geneseq/geneseq/geneseq/geneseq/AA1985.DAT:\*

7: /SIDS1/gcdata/geneseq/geneseq/geneseq/geneseq/AA1986.DAT:\*

8: /SIDS1/gcdata/geneseq/geneseq/geneseq/geneseq/AA1987.DAT:\*

9: /SIDS1/gcdata/geneseq/geneseq/geneseq/geneseq/AA1988.DAT:\*

10: /SIDS1/gcdata/geneseq/geneseq/geneseq/geneseq/AA1989.DAT:\*

11: /SIDS1/gcdata/geneseq/geneseq/geneseq/geneseq/AA1990.DAT:\*

12: /SIDS1/gcdata/geneseq/geneseq/geneseq/geneseq/AA1991.DAT:\*

13: /SIDS1/gcdata/geneseq/geneseq/geneseq/geneseq/AA1992.DAT:\*

14: /SIDS1/gcdata/geneseq/geneseq/geneseq/geneseq/AA1993.DAT:\*

15: /SIDS1/gcdata/geneseq/geneseq/geneseq/geneseq/AA1994.DAT:\*

16: /SIDS1/gcdata/geneseq/geneseq/geneseq/geneseq/AA1995.DAT:\*

17: /SIDS1/gcdata/geneseq/geneseq/geneseq/geneseq/AA1996.DAT:\*

18: /SIDS1/gcdata/geneseq/geneseq/geneseq/geneseq/AA1997.DAT:\*

19: /SIDS1/gcdata/geneseq/geneseq/geneseq/geneseq/AA1998.DAT:\*

20: /SIDS1/gcdata/geneseq/geneseq/geneseq/geneseq/AA1999.DAT:\*

21: /SIDS1/gcdata/geneseq/geneseq/geneseq/geneseq/AA2000.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	250	100.0	666	19 W62829 Macadamia integrifolia
2	242	96.8	666	19 W62828 Macadamia integrifolia
3	235	94.0	625	19 W62830 Macadamia integrifolia
4	112	44.8	525	19 W62831 Theobroma cacao an
5	112	44.8	566	13 R20181 Sequence encoded by
6	104	41.6	590	19 W62832 Gossypium hirsutum
7	87	34.8	28	19 W62841 Stenocarpus sinuatus
8	66.5	26.6	342	20 Y16785 Human secreted protein
9	66.5	26.6	637	19 W62837 Hordeum vulgare an
10	64.5	25.8	35	13 R21079 Antimicrobial maize
11	64.5	25.8	593	19 W62835 Zea mays antimicrobial
12	63.5	33	19	W62836 Zea mays antimicrobial

## ALIGNMENTS

RESULT	1	W62829	standard: protein; 666 AA.
ID	W62829		
XX			
AC	W62829;		
XX			
DT	27-OCT-1998 (first entry)		
XX			
DE	Macadamia integrifolia antimicrobial protein.		
XX			
KW	antimicrobial protein; infestation; control.		
XX			
OS	Macadamia integrifolia.		
XX			
PH	Key peptide	Location/Qualifiers	
FT		/note= "signal peptide"	
FT	Protein	1..28 /note= "mature protein"	
XX			
PN	W09827805-A1.		
XX			
PD	02-JUL-1998.		
XX			
PF	22-DEC-1997;	97WO-AU00874.	
XX			
PR	20-DEC-1996;	96AU-0004275.	
XX			
(RETR-)	COOP RES CENT TROPICAL PLANT PATHOLOGY.		
PA			
PT	Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;		
XX			
WPI:	1998-37729/32.		
DR	N-PSDD; VA2311.		
XX			

PT Novel anti-microbial protein from e.g. Macadamia integrifolia -  
 PT useful for controlling microbial infestations of plants or mammals  
 XX  
 PS Claim 1; Page 39-41; 96pp; English.

CC The sequence is that of an antimicrobial protein which can  
 CC be used to control microbial infestations in plants and mammalian  
 XX animals.  
 SQ Sequence 666 AA;

Query Match 100.0%; Score 230; DB 19; Length 666;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-20;  
 Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NQDPDPTDCQCQRCRQESGPRQQYCQRCKEICEEEFY 43  
 Db 74 nqddpqtdcgcqcrqrqgsgprqqycqrckeiceeeey 116

RESULT 2  
 ID W62828 standard; Protein: 666 AA.  
 AC W62828;  
 XX  
 DT 27-OCT-1998 (first entry)  
 DE Macadamia integrifolia antimicrobial protein.  
 XX  
 KW antimicrobial protein; infestation; control.  
 XX  
 OS Macadamia integrifolia.  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 1.28  
 FT /note= "signal peptide"  
 FT Protein 29..666  
 FT /note= "mature protein"  
 XX  
 PN W09827805-A1.  
 XX  
 PD 02-JUL-1998.  
 XX  
 PR 22-DEC-1997; 97WO-AU00874.  
 XX  
 BR 20-DEC-1996; 96AU-0004275.  
 XX  
 PA (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.  
 XX  
 PT Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;  
 XX  
 DR WPI; 1998-377279/32.  
 XX  
 N-PSDB; V42316.  
 XX  
 PT Novel anti-microbial protein from e.g. Macadamia integrifolia -  
 PT useful for controlling microbial infestations of plants or mammals  
 XX  
 PS Claim 1; Page 43-45; 96pp; English.  
 XX  
 The sequence is that of an antimicrobial protein which can  
 CC be used to control microbial infestations in plants and mammalian  
 CC animals.  
 XX  
 SQ Sequence 625 AA;

Query Match 94.0%; Score 235; DB 19; Length 625;  
 Best Local Similarity 93.0%; Pred. No. 6.9e-19;  
 Matches 40; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NQDPDPTDCQCQRCRQESGPRQQYCQRCKEICEEEFY 43  
 Db 33 nqedpqtdcgcqcrqrqgsgprqqycqrckeiceeeey 75

RESULT 4  
 ID W62831 standard; Protein: 525 AA.  
 AC W62831;  
 XX  
 DT 27-OCT-1998 (first entry)  
 XX  
 DE Theobroma cacao antimicrobial protein.  
 XX  
 KW antimicrobial protein; infestation; control.  
 XX  
 OS Theobroma cacao.

Query Match 96.8%; Score 242; DB 19; Length 666;  
 Best Local Similarity 95.3%; Pred. No. 1.2e-19;  
 Matches 41; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 NQDPDPTDCQCQRCRQESGPRQQYCQRCKEICEEEFY 43

XX WO9827805-A1.  
 XX PN RNA  
 XX CC detected in a cDNA library prepared from immature cocoa beans RNA  
 PD 02-JUL-1998.  
 XX CC using a probe based on the AA sequence of a CBR peptide common to  
 XX CC the 47 kD and 31 kD polypeptides. Homology searches revealed close  
 PT homologies between the 67 kD polypeptide and the vicilins, which are  
 PF seed storage proteins.  
 XX SQ Sequence 566 AA:  
 PR 20-DEC-1996; 96AU-0004275.  
 PA (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.  
 PI Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;  
 XX DR WPI: 1998-377279/32.  
 XX Novel anti-microbial protein from e.g. Macadamia integrifolia -  
 PT useful for controlling microbial infestations of plants or mammals  
 XX PS Claim 1; Page 47-49; 96PP; English.  
 XX The sequence is that of an antimicrobial protein which can  
 CC be used to control microbial infestations in plants and mammalian  
 CC animals.  
 XX SQ Sequence 525 AA:  
 RESULT 5  
 R20181 Query Match 44.8%; Score 112; DB 19; Length 525;  
 ID R20181 Best Local Similarity 47.5%; Pred. No. 3.1e-05;  
 DE Matches 19; Conservative 10; Mismatches 11; Indels 0; Gaps 0;  
 AC R20181  
 QY 2 QDPDPOTCQCCORRCROQESGRQQYCQRCKELEEEE 41  
 DB 78 eeelqrqyqqcqgrcqeqgggqregqgcqrkcdwykeqe 117  
 XX SQ Sequence 566 AA:  
 RESULT 6  
 W62832 Query Match 44.8%; Score 112; DB 13; Length 566;  
 ID W62832 Best Local Similarity 47.5%; Pred. No. 3.4e-05;  
 DE Matches 19; Conservative 10; Mismatches 11; Indels 0; Gaps 0;  
 AC W62832;  
 XX DT 27-OCT-1998 (first entry)  
 XX DE Gossypium hirsutum antimicrobial protein.  
 XX KW antimiicrobial protein; infestation; control.  
 XX OS Gossypium hirsutum.  
 XX PN WO9827805-A1.  
 XX PD 02-JUL-1998.  
 XX PR 22-DEC-1997; 97WO-AU00874.  
 XX PR 20-DEC-1996; 96AU-0004275.  
 AC R20181  
 XX SQ Sequence 590 AA:  
 RESULT 5  
 DT 16-APR-1992 (first entry)  
 DE Sequence encoded by 67 kD T. cacao protein cDNA.  
 KW Cocoa; flavour; vicilin; seed storage protein.  
 XX OS Theobroma cacao.  
 XX PI Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;  
 DE WPI: 1998-377279/32.  
 XX PR (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.  
 XX PT Novel anti-microbial protein from e.g. Macadamia integrifolia -  
 XX PT useful for controlling microbial infestations of plants or mammals  
 XX PS Claim 1; Page 49-51; 96PP; English.  
 XX The sequence is that of an antimicrobial protein which can  
 CC be used to control microbial infestations in plants and mammalian  
 CC animals.  
 XX SQ Sequence 590 AA:  
 RESULT 7  
 W62841 Query Match 41.6%; Score 104; DB 19; Length 590;  
 ID W62841 Best Local Similarity 47.5%; Pred. No. 0.00027;  
 DE Matches 19; Conservative 9; Mismatches 10; Indels 2; Gaps 1;  
 AC W62841  
 QY 3 DDPDPOTCQCCORRCROQESGRQQYCQRCKELEEEE 42  
 DB 81 edpqqrqyeeeqcqcqeq -rqqpqcqcqrkckfeqeqq 118  
 XX SQ Sequence 590 AA:  
 RESULT 7  
 W62841 The inventors claim a 67 kD and 31 kD T. cacao protein, and  
 XX fragments, and encoding DNAs. The 47 kD and 31 kD proteins are  
 CC

XX  
 DT 27-OCT-1998 (first entry)  
 XX DE Stenocarpus sinuatus antimicrobial protein.  
 XX KW antimicrobial protein; infestation; control.  
 XX OS Stenocarpus sinuatus.  
 XX PN WO9827805-A1.  
 XX PD 02-JUL-1998.  
 XX PR 22-DEC-1997; 97WO-AU00874.  
 XX PR 20-DEC-1996; 96AU-0004275.  
 XX PA (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.  
 XX PI Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;  
 XX DR WPI; 1998-377279/32.  
 XX PT Novel anti-microbial protein from e.g. Macadamia integrifolia - useful for controlling microbial infestations of plants or mammals  
 XX PS Claim 1; Page 66; 96pp; English.  
 CC The sequence is that of an antimicrobial protein which can be used to control microbial infestations in plants and mammalian animals.  
 XX SQ Sequence 28 AA;  
 Query Match 34.8%; Score 87; DB 19; Length 28;  
 Best Local Similarity 59.3%; Pred. No. 0.0012; Matches 16; Conservative 2; Mismatches 9; Indels 0; Gaps 0;  
 QY . 4 DPQTCQCQCORRCROQESPRQQYCQ 30  
 Db 2 dpircqqlcamrcgqekprqqgck 28

RESULT 8  
 Y16785 ID Y16785 standard; Protein: 342 AA.  
 XX AC Y16785;  
 XX DT 27-JUL-1999 (first entry)  
 XX DE Human secreted protein (clone ct489\_14).  
 XX KW Secreted protein; human; tissue marker; genetic disease; gene therapy; veterinary medicine; cell proliferation; immunostimulant; infection; immunosuppressant; autoimmune disease; organ rejection; tumour; anaemia; haematoopoiesis; wound healing; fertility control; chemotaxis; analgesic; thrombolytic; haemophilia; infarction; antimicrobial agent; cancer.  
 XX OS Homo sapiens.  
 XX PN WO9924469-A1.  
 XX PD 20-MAY-1999.  
 XX PF 06-NOV-1998; 98WO-US23829.  
 XX PR 04-NOV-1998; 98US-0105936.  
 XX PR 07-NOV-1997; 97US-0965789.  
 XX PA (GENY ) GENETICS INST INC.

RESULT 9  
 W62837 ID W62837 standard; Protein: 637 AA.  
 XX AC W62837;  
 XX DT 27-OCT-1998 (first entry)  
 XX DE Hordeum vulgare antimicrobial protein.  
 XX KW antimicrobial protein; infestation; control.  
 XX OS Hordeum vulgare.  
 XX PN WO9927805-A1.  
 XX PD 02-JUL-1998.  
 XX PF 22-DEC-1997; 97WO-AU00874.  
 XX PR 20-DEC-1996; 96AU-0004275.  
 XX PA (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.  
 XX PI Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;  
 XX DR WPI; 1998-377279/32.  
 XX PT Novel anti-microbial protein from e.g. Macadamia integrifolia - useful for controlling microbial infestations of plants or mammals  
 XX



XX  
 DR WPI; 1998-377279/32.  
 XX Novel anti-microbial protein from e.g. *Macadamia integrifolia* -  
 PT useful for controlling microbial infestations of plants or mammals  
 XX  
 PS Disclosure; Page 60; 96pp; English.  
 XX  
 CC The sequence is that of an antimicrobial protein which can  
 CC be used to control microbial infestations in plants and mammalian  
 CC animals.  
 XX  
 SQ Sequence 33 AA:  
 Query Match 25.4%; Score 63.5; DB 19; Length 33;  
 Best Local Similarity 44.0%; Pred. No. 0.6;  
 Matches 11; Conservative 6; Mismatches 7; Indels 1; Gaps 1;  
 Qy 11 OCQRRC-RQOESGPRQQYCQRRCK 34  
 ID :|:||: |: | ; | ||:|  
 Db 6 ecrrqclrrheggpwetqecmrcc 30  
 RESULT 13  
 R91705  
 ID R91705 standard; Protein; 107 AA.  
 XX  
 AC R91705;  
 XX  
 DT 17-NOV-1996 (first entry)  
 XX  
 DE AcanAP23.  
 XX  
 KW AcANAP; HpoNAP; NamNAP; AcenNAP; AdUNAP; anticoagulant;  
 KW nematode-extracted anticoagulant protein; serine protease;  
 KW nematode; thrombosis; parasitic worm.  
 OS Ancylostoma caninum.  
 XX  
 PN W09612021-A2.  
 XX  
 PD 25-APR-1996.  
 XX  
 PR 17-OCT-1995; 95WO-US13231.  
 XX  
 PR 05-JUN-1995; 95US-0486399.  
 PR 18-OCT-1994; 94US-0326110.  
 PR 05-JUN-1995; 95US-0463965.  
 PR 05-JUN-1995; 95US-0463380.  
 PR 05-JUN-1995; 95US-0486397.  
 PA (CORV-) CORVAS INT INC.  
 XX  
 PT Bergum PW, Ganssmans YGJ, Jaspers LS, Laroche YR;  
 PT Lauwereys MJ, Messens JHL, Moyle M, Stanssens PEH;  
 PI Vlasuk GP;  
 XX  
 DR WPI; 1996-222007/22.  
 DR N-PSDB; T12951.  
 XX  
 PT Proteins with anticoagulant and/or serine protease inhibitory  
 PT activity, isolated from nematodes and useful to inhibit blood  
 coagulation.  
 PT The proteins can be added to blood collection tubes  
 defining the collection of mammalian plasma. They are also useful  
 to prevent or inhibit thrombosis, and may be given alone or in  
 combination with other therapeutic or in vivo diagnostic agents.  
 XX  
 CC The proteins can serve as immunogens to raise antibodies for use in  
 CC the diagnosis and identification of NAP concn. Levels in biological  
 CC fluids, e.g. to detect mammalian infection with a parasitic worm.  
 CC They can also be used as immunogens in prophylactic and therapeutic  
 CC vaccines against parasitic worm infection. The proteins may  
 CC double the clotting time of human plasma in prothrombin time assays  
 CC when present at 10-50 nMOL.  
 CC Plasma in activated partial thrombin time assays when present  
 CC at 10-100 nMOL.  
 CC The anticoagulant proteins are pref. derived from  
 CC *Ancylostoma caninum*, *A. ceylanicum*, *A. duodenale*, *Necator  
 americanus* or *Heligmosomoides polygyrus*.  
 CC The proteins pref. have 2 NAP domains and specifically inhibit  
 CC the catalytic activity of the factor VIIa/TF complex in the  
 CC presence of Factor Xa or a catalytically inactive factor Xa deriv.,  
 CC do not specifically inhibit the activation of factor VIIa in the  
 CC absence of TF and do not specifically inhibit prothrombinase.  
 XX  
 SQ Sequence 107 AA;  
 Query Match 25.0%; Score 62.5; DB 17; Length 107;  
 Best Local Similarity 33.3%; Pred. No. 2.3;  
 Matches 13; Conservative 10; Mismatches 9; Indels 7; Gaps 2;  
 Qy 12 CQRCRQOESGPRO-----QYQCORRCK-EICBEEEY 43  
 ID | : ::| | : |:|:| | ||:|  
 Db 17 ctgkpsekecgpherdcgknkkpcerkkietseeddy 55  
 RESULT 14  
 Y30404  
 ID Y30404 standard; Protein; 107 AA.  
 XX  
 AC Y30404;  
 XX  
 DT 15-NOV-1999 (first entry)  
 XX  
 DE Nematode extracted anticoagulant protein AcanAP23.  
 XX  
 KW Nematode extracted anticoagulant protein; NAP; anticoagulant;  
 KW serine protease inhibitor; NAP domain; factor VIIa/TF.  
 XX  
 OS Ancylostoma caninum.  
 XX  
 PN US955294-A.  
 XX  
 PD 21-SEP-1999.  
 XX  
 PR 19-APR-1996; 96US-0634641.  
 PR 19-APR-1996; 96US-0634641.  
 PR 18-OCT-1994; 94US-0326110.  
 PR 05-JUN-1995; 95US-0463965.  
 PR 05-JUN-1995; 95US-0463380.  
 PR 05-JUN-1995; 95US-0486397.  
 PR 05-JUN-1995; 95US-0486397.  
 PR 17-OCT-1995; 95WO-US13231.  
 XX  
 PA (CORV-) CORVAS INT INC.  
 XX  
 PT Bergum PW, Ganssmans YGJ, Jaspers LS, Laroche YR;  
 PT Lauwereys MJ, Messens JHL, Moyle M, Stanssens PEH;  
 PI Vlasuk GP;  
 XX  
 DR WPI; 1999-539569/45.  
 DR N-PSDB; Z10452.  
 XX  
 PT Screening an isolated protein for Nematode-extracted Anticoagulant  
 PT Protein domains  
 XX  
 PS Example 12; Fig 13A; 197pp; English.  
 XX



